Homework 3

Christophe Hunt

February 18, 2017

Contents

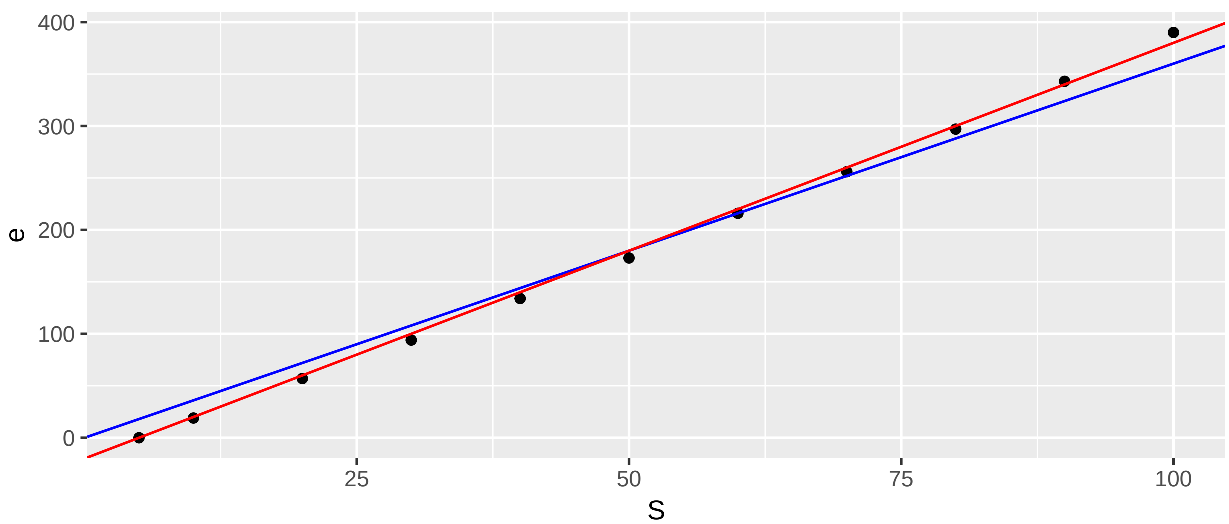
1 Problem : Page 113: 2

The following table gives the elongation e in inches (in./in.) for a given stress S on a steel wire measured in pounds per square inch (lb/in.^{2}). Test the models $e = c_1 S$ by plotting the data. Estimate c_1 graphically.

Table 1:

$S(x10^{-3})$	5	10	20	30	40	50	60	70	80	90	100
$e(x10^5)$	0	19	57	94	134	173	216	256	297	343	390

```r
library(ggplot2)
S <- c(5,10,20,30,40,50,60,70,80,90,100)
e <- c(0,19,57,94,134,173,216,256,297,343,390)
ggplot(data = as.data.frame(cbind(S,e)), aes(x = S, y = e)) +
  geom_point() +
  geom_abline(slope = 3.6, color = 'blue') +
  geom_abline(intercept = -20, slope = 4, color = 'red')
```



Above is the graph of the elongation $e% versus stress S x 10^{-1}. By eyeballing the results of several plots we can give the estimate of ~3.6 for c_1 for the model $e = c_1 S$ (this is the blue line). However, do see a much better fit with ~4 for c_1, if we provide an intercept of -20. These are simply best guesses.

2 Problem : Page 121: 2.a

For each of the following data sets, formulate the mathematical model that minimizes the largest deviation between the data and the line y = ax + b. If a computer is available, solve for the estimates of a and b.

Table 2:

x	1	2.3	3.7	4.2	6.1	7.0
y	3.6	3.0	3.2	5.1	5.3	6.8

```
x <- c(1  , 2.3, 3.7, 4.2, 6.1, 7.0)
y <- c(3.6, 3.0, 3.2, 5.1, 5.3, 6.8)

mean.x <- mean(x)
mean.y <- mean(y)

x.i <- (x - mean.x)
y.i <- (y - mean.y)


x.i.y.i <- (x.i * y.i)
x.i.2   <- (x.i^2)

m <- sum(x.i.y.i) / sum(x.i.2)
b <- mean.y - m*mean.x

y2 <- (m*x + b)
d.max <- max(abs(y - y2))
```

The model y = ax + b for this data is $y = 0.56x + 2.21$, with a $d_{max} = 1.1025182$.

3 Problem : Page 127: 10

Data For planets

Body	Period (sec)	Distance from sun (m)
Mercury	7.60 x 10^6	5.79 x 10^10
Venus	1.94 x 10^7	1.08 x 10^11
Earth	3.16 x 10^7	1.5 x 10^11
Mars	5.94 x 10^7	2.28 x 10^11
Jupiter	3.74 x 10^8	7.79 x 10^11
Saturn	9.35 x 10^8	1.43 x 10^12
Uranus	2.64 x 10^9	2.87 x 10^12
Neptune	5.22 x 10^9	4.5 x 10^12

Fit the model $y = ax^{3/2}$

```
period <- c(( 7.60 * 10^6 ), ( 1.94 * 10^7 ), ( 3.16 * 10^7 ),
            ( 5.94 * 10^7 ), ( 3.74 * 10^8 ), ( 9.35 * 10^8 ),
            ( 2.64 * 10^9 ), ( 5.22 * 10^9 ))

distances <- c(( 5.79 * 10^10 ), ( 1.08 * 10^11 ), ( 1.5 * 10^11  ),
               ( 2.28 * 10^11 ), ( 7.79 * 10^11 ), ( 1.43 * 10^12 ),
               ( 2.87 * 10^12 ), ( 4.5 * 10^12  ))
```

Least square solution to the formula $y = An^x$, for the model $y = an^{3/2}$.

```
a <- sum(period^(3/2) * distances)/sum((period^2)^(3/2))
a
```

```
## [1] 0.01320756
```

Resulting in the form $y = 0.0132n^{3/2}$.

4 Problem : Page 136: 7

4.1 a.

In the following data, W represents the weight of a fish (bass) and l represents its length. Fit the model $W = kl^3$ to the data using the least-squares criterion.

Length, l (in.)	14.5	12.5	17.25	14.5	12.625	17.75	14.125	12.635
Weight, W (oz)	27	17	41	26	17	49	23	16

```
x <- length.in <- c(14.5, 12.5, 17.25 , 14.5 , 12.625 , 17.75 , 14.125 , 12.625)
y <- weight.oz <- c(27 , 17    , 41    , 26   , 17     , 49    , 23     , 16 )

a <- sum(x^3*y)/(sum((x^2)^3))
y2 <- a*(x^3)
y.y2 <- (y - y2)
D <- sqrt(sum(y.y2^2)/8)
```

The least-squares fit of $W = kl^3$ is $W = 0.008437l^3$. The sum of the squares of the deviations as 12.1683418 so $D = 1.2333056$. As the largest absoulte deviation is 2.305, c_{max} can be bound as follows: $D = 1.2333056 \leq c_{max} \leq 2.305 = d_{max}$

4.2 b.

In the following data, g represents the girth of a fish. Fit the model $W = klg^2$ to the data using the least squares criterion

Length, l (in.)	14.5	12.5	17.25	14.5	12.625	17.75	14.125	12.625
Girth g (in)	9.75	8.375	11	9.75	8.5	12.5	9.0	8.5
Weight, W (oz)	27	17	41	26	17	49	23	16

```
x <- length.in <- c(14.5, 12.5 , 17.25, 14.5, 12.625, 17.75 , 14.125 , 12.625)
y <- weight.oz <- c(27   , 17   , 41    , 26   , 17    , 49    , 23     , 16 )
z <- girth.in  <- c(9.75, 8.375, 11    , 9.75, 8.5   , 12.5  , 9.0    , 8.5)

a <- sum((x*z^2)*y)/(sum((x*z^2)^2))
y2 <- a*(x*z^2)
y.y2 <- (y - y2)
D <- sqrt(sum(y.y2^2)/8)
```

The least-squares fit of $W = klg^2$ is $W = 0.018675lg^2$. The sum of the squares of the deviations as 17.6710973 so $D = 1.4862325$. As the largest absoulte deviation is 2.794, c_{max} can be bound as follows: $D = 1.4862325 \leq c_{max} \leq 2.794 = d_{max}$

4.3 c.

Which of the two models fits the data better? Justify fully. Which model do you prefer? Why?

The first model has the largest absolute deviation at 2.305, whereas, the second model's largest absolute deviation at 2.794. Therefore, it appears that the first model fits the data better with the lowest d_{max}. However, using only the criterion may be naive for the data set. My preference would be for the second model because of its account for girth which would suggest a growth in weight and has the smallest d_{max} of the two models.